Sean Yang

Master Data Management Specialist

xiaoyang_yang@outlook.com · Shenzhen, China · yxyang.info · github.com/toadish-frog · www.linkedin.com/in/xya89/

SUMMARY

Data analyst and tooling engineer with 3+ years across supply chain, catalog, and operations. I build automation and BI systems that replace manual, error-prone workflows with something a team can run itself, using agentic coding tools (Gemini CLI, Claude Code) as a core part of how I design and ship them. At Thrasio, I own the NetSuite catalog data model, cut SKU processing SLAs by 95% with Python automation, and built the team's Power BI and Looker Studio reporting stack. Core stack: Python, SQL, NetSuite, GCP, Power BI, Claude Code.

EXPERIENCE

Master Data Management Specialist

Feb 2025 – Present

Thrasio · Shenzhen, China

- Rebuilt the SKU creation and update workflow as a full-stack Streamlit app built with Gemini CLI as an AI coding agent, cutting the average SLA from 24 hours to 30 minutes (a 95% reduction) and reducing a two-person job to one person.
- Automated multi-step data validation and NetSuite integration, eliminating an estimated 5 hours of manual spreadsheet work per analyst per week.
- Implemented Google OAuth authentication for internal tooling, enabling secure, role-based access for MDM team members.
- Led the rebuild of NetSuite's item and channel data model, using Gemini CLI to reverse-engineer undocumented legacy tooling before restoring a single source of truth for catalog data.
- Authored and rewrote 3 MDM SOPs (SKU creation/update, dimension & weight validation, 3PL master lookup), taking dimension/weight validation to 100% coverage across all 2025 outbound POs.
- Built the team's Power BI and Looker Studio KPI dashboards (SKU maintenance, Amazon Order-to-Cash, Stord monthly transactions) and led the team's migration from Power BI to Looker Studio.
- Ran internal training sessions on Google Apps Script automation, Looker Studio, and GitHub that helped teammates build their own data pipelines and dashboards, cutting out a recurring manual download/upload task.

Business Analyst

May 2023 – Feb 2024

Norsat International Inc. · Richmond, Canada

- Built Power BI dashboards for the purchasing team covering Inventory Stocking Levels and Purchase Order Tracking, using Azure SQL and Business Central Web Services as data sources.
- Reduced report refresh time by 50% by standardizing data fetching methods and query patterns across the BI layer.
- Maintained a 95% on-time delivery rate for purchase orders by coordinating across marketing, planning, and warehouse teams.
- Identified and corrected KPI calculation errors in presale and supply chain reports, contributing to approximately $70,000 in annual cost savings.

Operations Assistant (Intern)

Jan 2023 – Apr 2023

Alba Partners Inc. · Vancouver, BC

- Built operational and sales reports in Excel and ERP, reducing manual lookup time.
- Modeled company-specific trends using SQL, Excel, and Python on internal datasets, improving data processing efficiency by 10%.

EDUCATION

B.Sc. Data Science
Sep 2018 – Apr 2023

Simon Fraser University · Burnaby, Canada

- Dean's List 2021
- Vancouver Coastal Health: Canines-for-Care (CDI forecasting with VGH data analysts)
- Steam Server Usage Forecast (seasonal ARIMA time-series study)

PROJECTS

Agent Andon Light
Jul 2026 – Present

Personal

- Physical Andon-style status light for coding-agent sessions, driven automatically by hook events in both Claude Code and Codex CLI, so I can walk away from the terminal and still know at a glance whether the agent is working, waiting, or stuck. Grew from a 3-LED breadboard prototype into a custom PCB carrying 10 individually addressable LEDs, a 3D-printed enclosure, and a PyPI-published host CLI with a one-click Windows installer, alongside a documentation and marketing site translated into 4 languages.
- Designed a 5-state wire protocol and firmware watchdog (RP2040/C++) driving 10 individually addressable LEDs across a custom PCB, failing safe to a distinct "stale" color instead of silently freezing if the host driver ever crashes.
- Found and fixed a race condition caused by async hook execution, where a delayed background command could silently overwrite a later, correct state: a bug only visible under real session load.
- Extended hook integration from Claude Code to a second agent, Codex CLI, as a host-only change against the same wire protocol and firmware: no reflash needed to support a second agent.
- Took the project from a breadboard MVP through a custom PCB and 3D-printed enclosure: 5 units built, assembled, and verified end-to-end on real hardware.
- Packaged and published the host CLI to PyPI (`andon-light`) with a one-click Windows installer (PyInstaller + Inno Setup), cutting install to two commands on Linux and a guided GUI flow on Windows.
- Designed and built a companion documentation and marketing site (Next.js, static export) covering install guides, hook references, and troubleshooting docs, localized across English, Japanese, Simplified Chinese, and Traditional Chinese.

Automation Olympics: SKU Request Processor
May 2025 – Present

Thrasio

- Streamlit web app that takes a SKU update or creation request from validation through NetSuite updates to email notification in a single workflow, replacing a fragmented multi-step spreadsheet process.
- Architected and built the app with Gemini CLI as an AI coding agent, covering design, implementation, and code review, learning firsthand where agentic coding needs human oversight.
- Won 1st place in Thrasio's Automation Olympics, an internal competition for team-built automation tools.

Vancouver Coastal Health: Canines-for-Care
Jan 2023 – Apr 2023

Vancouver Coastal Health / Simon Fraser University

- Collaborated with data analysts from Vancouver General Hospital to develop predictive clustering models forecasting CDI outbreaks from canine screening data collected between 2017 and 2022.
- Built CDI forecasting models in R using clustering and correlation analysis on 5 years of canine screening and infection records.
- Delivered exploratory and risk analysis in Tableau, surfacing correlations between canine alert frequency and CDI case counts.

Steam Server Utilization Forecasting
Nov 2022 – Dec 2022

Simon Fraser University

- Used the Box-Jenkins method and dynamic regression approaches on SteamDB data (2020–2022) to fit a multiplicative seasonal ARIMA model and produce a 6-month server utilization forecast.
- Fit a seasonal ARIMA(2,1,2) × (0,1,1) model with a 7-day period using R.
- Forecast 9,487,428 peak concurrent players against an actual of 9,484,341, a 99.97% accuracy rate.

SKILLS

Languages:	Python · SQL · TypeScript · Google Apps Script · C · C++ · R
AI-Native Development:	Agentic Coding Tools (Gemini CLI, Claude Code) · AI-Assisted Systems Architecture & Code Review · Rapid AI-Assisted Prototyping ("Vibe Coding") · Legacy System Reverse-Engineering with AI · Full-Stack AI-Assisted Builds (Hardware to Software)
Frameworks & Libraries:	Streamlit · Next.js · React
Tools & Platforms:	Google Cloud Platform · Oracle NetSuite · Power BI · Looker Studio · Snowflake · dbt · Azure SQL · Docker · Claude Code · CI/CD (GitHub Actions) · Arduino / Embedded Systems (RP2040) · Git · Draw.io · Zendesk
Soft Skills:	Master Data Management · Data Modeling · Supply Chain Operations · Cross-functional Collaboration · Technical Mentorship · Technical Writing · Process Documentation